

# On Composite Discontinuous Galerkin Method for simulations of electric properties of semiconductor devices

Konrad Sakowski<sup>\*1,3</sup>, Leszek Marcinkowski<sup>3</sup>, Pawel Strak<sup>1</sup>, Pawel Kempisty<sup>1</sup>, and Stanislaw Krukowski<sup>1,2</sup>

<sup>1</sup>Institute of High Pressure Physics, Polish Academy of Sciences,  
ul. Sokolowska 29/37, 01-142 Warsaw, Poland

<sup>2</sup>Interdisciplinary Centre for Materials Modeling, Warsaw  
University, ul. Pawińskiego 5a, 02-106 Warsaw, Poland

<sup>3</sup>Department of Mathematics, Computer Science and Mechanics,  
Warsaw University, ul. Banacha 2, 02-097 Warsaw, Poland

March 4, 2016

## Abstract

In this paper two variants of discretization of the van Roosbroeck's equations in equilibrium state with the Composite Discontinuous Galerkin Method for rectangular domain are discussed. They base on Weakly Over-Penalized Symmetric Interior Penalty (WOPSIP) method and on Symmetric Interior Penalty Galerkin (SIPG) method. It is shown that the discrete problems are well-defined and that their solutions are unique. Error estimates are derived. Finally numerical simulations of gallium nitride semiconductor devices are presented.

## 1 Introduction

Numerical simulations are the important tool in development of semiconductor devices. Since our contemporary electronics relies on the semiconductors, there is strong demand on the progress in this domain. Examples of such devices are light emitting diodes, lasers, transistors, detectors, and many others. There are various approaches in simulations of such devices, depending on precision, efficiency and size of a simulated fragment. On the one hand there are so-called *ab initio* methods, which are used to investigate properties of elements composed of hundreds of thousands of atoms. These methods use fundamental laws of physics and they need days or weeks to perform a single simulation on a computational cluster. On the other hand there is a drift-diffusion theory.

---

<sup>\*</sup>konrad@unipress.waw.pl

In this case the model is much simpler and it allows to simulate whole semiconductor device on a standard desktop computer. This model describes two kinds of carriers, electrons and holes, which move in the electric field present in semiconductor devices. From the mathematical point of view, it consists of a system of three nonlinear elliptic differential equations, which are called van Roosbroeck equations [14].

Numerical modelling of semiconductor devices with the drift-diffusion model are performed since 1964, when Gummel [8] proposed a numerical algorithm based on the simple iteration method. Various methods were used for discretization of the van Roosbroeck equations, for example Finite Difference Method [17], Box method [1], Finite Element Method (FEM) [6]. Special variants of discretizations optimized for the so-called continuity equations were developed [11].

In our study, we are interested in laser and electroluminescent diodes based on gallium nitride. These devices have particular structure. They are composed of layers deposited one on another. These layers vary in material composition, doping level, dopant type or length, and they form a natural partition of a device. They are usually cuboids (see figure 1). In general, in these devices there are three important parts: two contacts and the active region. A potential can be applied to the contacts and then some current will flow through the device. Then, in the active region, the recombination occurs and fotons are generated.

Due to such a structure, it is often sufficient to use one-dimensional model for simulations of the laser diodes and the electroluminescent diodes, because unless a device has some special design, the current flows more or less in the straight line between the contacts. However, to take into account more sophisticated designs, more spatial dimensions have to be used.

Motivated by this problem, in this paper we study the Composite Discontinuous Galerkin Method discretizations. In these methods, first we pick some partition  $\{\Omega_i\}_{i=1}^N$  of a domain  $\Omega$ . Then on every element  $\Omega_i$  we introduce some triangulation and a Finite Element space  $X_{h_i}(\Omega_i)$ . These triangulations may be independent of each other, i.e. they do not need to agree on interfaces  $\partial\Omega_i \cap \partial\Omega_j$ . Finally we search for a discrete solution, which is in the space  $X_{h_i}(\Omega_i)$  on every  $\Omega_i$ . On the interfaces we use a variant of the Discontinuous Galerkin Method (DGM) to “glue” functions from the different partition elements together.

Our interest in application of these methods to modelling of the semiconductor luminescent devices is caused by the following reasons. First, the natural division of these devices described above is a good candidate for a partition of a computational domain  $\Omega$ . It is so because the discontinuities of the coefficients are often on interfaces between layers due to variations of material parameters. Also it is often the case that one could easily predict on which layers the variations of the unknown functions are more frequent than on the others, due to physical properties of the device. Therefore it convenient to use separate meshes, as there is no problem in mesh refinement in arbitrary layers.

We start with introduction of the differential problem in section 2. We propose two variants of Composite DGM discretization of this problem in section 3. Then we show existence and uniqueness of the introduced discrete problems in sections 4, 5. In section 6 we discuss interpolation properties of the discrete space. Then we pass to the error estimates in section 7. Finally we present results of numerical simulations in section 8 and we conclude in section 9.

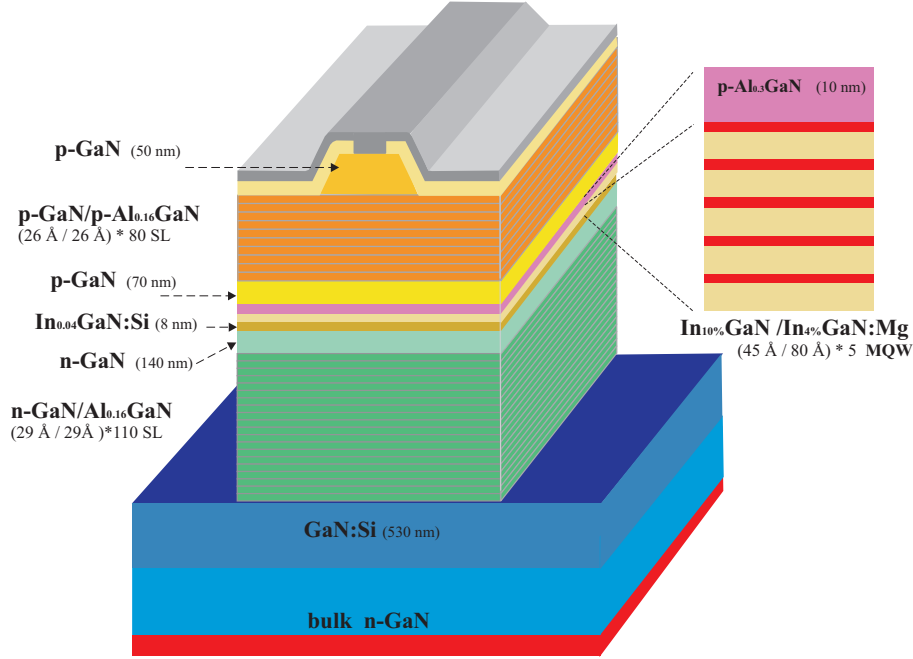


Figure 1: Example of a gallium nitride semiconductor laser structure.

## 2 Differential problem

The drift-diffusion model describes relationship between the electrostatic potential and the charge carrier concentrations: electrons and holes [19, 18]. Physical derivation of this model is beyond the scope of this work, therefore we will focus on the mathematical standpoint.

We start with the domain  $\Omega$  of our problem. Luminescent semiconductor devices are generally made of planar layers deposited one on another, which vary in composition of a semiconductor material or number of impurities (see figure 1). At opposite ends metal contacts are attached, where the current can be applied. If this is the case, it flows through the device perpendicular to the deposited layers. We assume that  $\Omega$  is a rectangle.

In this paper we deal with so-called equilibrium state. It is a physical state of a semiconductor device, where the device is disconnected from power sources, so the current does not flow through it. Then only the Poisson equation needs to be solved and the system reduce to the following problem: find  $u^* \in H^1(\Omega)$ , such that

$$\begin{aligned} -\nabla \cdot (\varepsilon(x) \nabla u^*) + e^{u^* - \hat{v}} - e^{\hat{w} - u^*} &= k_1, \\ u^* &= \hat{u} \text{ on } \Sigma_D, \\ \nabla u^* \cdot \nu &= 0 \text{ on } \Sigma_N. \end{aligned} \tag{1}$$

where  $\hat{v} = \hat{w} \equiv \text{const.}$  Since some results of this paper may be also applied to non-equilibrium case, we consider more general assumption that  $\hat{v}, \hat{w} \in L_\infty(\Omega)$ . Also we assume that  $\varepsilon, \hat{u} \in H^1(\Omega) \cap L_\infty(\Omega)$  and  $0 < \varepsilon_m \leq \varepsilon \leq \varepsilon_M$ ,  $\varepsilon_m, \varepsilon_M \in \mathbb{R}$ .

$\Omega_7$	$\Omega_8$	$\Omega_9$	$\Omega_{10}$	$\Omega_{11}$	$\Omega_{12}$
$\Omega_1$	$\Omega_2$	$\Omega_3$	$\Omega_4$	$\Omega_5$	$\Omega_6$

Figure 2: An example of two-dimensional coarse grid of  $\Omega$ .

The following lemma is essential for the results presented in this paper. Its proof may be found in [9].

**Lemma 2.1.** *Solution  $u^*$  of problem (1) is bounded.*

A weak formulation of the differential problem (1) is as follows. Find  $u^* \in \hat{u} + H^1(\Omega)$ , such that

$$a(u^*, \varphi) + b(u^*, \varphi) = f(\varphi) \quad \forall \varphi \in H_{0,\Sigma_D}^1(\Omega), \quad (2)$$

where

$$\begin{aligned} a(u, \varphi) &:= \int_{\Omega} \varepsilon(x) \nabla u(x) \cdot \nabla \varphi(x) \, dx \\ b(u, \varphi) &:= \int_{\Omega} \left( e^{u(x)-\hat{v}(x)} - e^{\hat{w}(x)-u(x)} \right) \varphi(x) \, dx. \\ f(\varphi) &:= \int_{\Omega} k_1(x) \varphi(x) \, dx. \end{aligned} \quad (3)$$

We use the following notation:

$$\begin{aligned} C_{0,\Sigma_D}^\infty(\overline{\Omega}) &:= \{f \in C^\infty(\overline{\Omega}) : f|_{\Sigma_D} \equiv 0\}, \\ H_{0,\Sigma_D}^1(\Omega) &:= \text{cl}_{H^1(\Omega)} \left( C_{0,\Sigma_D}^\infty(\overline{\Omega}) \right). \end{aligned} \quad (4)$$

### 3 Discretization

#### 3.1 Discrete space

Let  $\Omega \subset \mathbb{R}^2$  be a rectangle, divided to disjoint subrectangles  $\{\Omega_i\}_{i=1}^N =: \mathcal{E}$  in such a manner that  $\mathcal{E}$  is a triangulation of  $\Omega$  (figure 2). We will call this division a coarse grid and we assume that each edge of any  $\Omega_i$  is contained either in  $\Sigma_D$  or in  $\Sigma_N$ .

Let us define triangulations  $\mathcal{T}_{h_i} := \mathcal{T}_{i,h_i}(\Omega_i)$ , where  $h_i := \max\{\text{diam}(\tau) : \tau \in \mathcal{T}_{h_i}\}$ . By  $\mathcal{N}_{h_i}$  we denote the nodes of the triangulation  $\mathcal{T}_{h_i}$ . We assume that  $\{\mathcal{T}_{i,h_i}(\Omega)\}_{h_i}$  is a regular uniform family of triangulations [5]. We will define

$\mathcal{T}_h := \bigcup_{i=1}^N \mathcal{T}_{h_i}$ . For  $s > 0$ , we define the broken Sobolev spaces  $H^s(\mathcal{E})$  and  $H^s(\mathcal{T}_h)$  as

$$\begin{aligned} H^s(\mathcal{E}) &:= \{v \in L_2(\Omega) : \forall i \in \{1, \dots, N\} \ v_i := v|_{\Omega_i} \in H^s(\Omega_i)\} \subset L_2(\Omega), \\ H^s(\mathcal{T}_h) &:= \{v \in L_2(\Omega) : \forall \tau \in \mathcal{T}_h \ v|_{\tau} \in H^s(\tau)\} \subset L_2(\Omega). \end{aligned} \quad (5)$$

Then on every  $\Omega_i$  we define a discrete space  $X_{h_i}(\Omega_i)$  of piecewise linear functions on the triangulation  $\mathcal{T}_{h_i}$ :

$$X_{h_i} := X_{h_i}(\Omega_i) := \left\{ u_{h,i} \in \mathcal{C}(\overline{\Omega_i}) : \forall \tau \in \mathcal{T}_{h_i} \ u_{h,i}|_{\tau} \in \mathbb{P}_1(\tau) \right\} \quad (6)$$

Finally we define  $X_h(\Omega)$  as

$$X_h(\Omega) = \left\{ (u_{h,1}, \dots, u_{h,N}) : u_{h,i} \in X_{h_i}(\Omega_i), i \in \{1, \dots, N\} \right\} \subset L_2(\Omega). \quad (7)$$

Note that  $X_h(\Omega) \not\subset H^1(\Omega)$  and  $X_h(\Omega) \not\subset H^2(\mathcal{E})$ , but  $X_h(\Omega) \subset H^1(\mathcal{E})$ ,  $H^1(\Omega) \subset H^1(\mathcal{E})$  and  $X_h(\Omega) \subset H^2(\mathcal{T}_h)$ .

By  $\Gamma$  we denote a set of all internal and boundary edges of  $\mathcal{E}$ . Then  $\Gamma$  is a sum of disjoint sets  $\Gamma_D$ ,  $\Gamma_N$  and  $\Gamma_I$ , where

$$\begin{aligned} \Gamma_D &:= \{e \in \Gamma : e \subset \Sigma_D\}, \\ \Gamma_N &:= \{e \in \Gamma : e \subset \Sigma_N\}, \\ \Gamma_I &:= \{e \in \Gamma : e \subset \text{int}(\Omega)\}. \end{aligned} \quad (8)$$

Therefore  $\Gamma_D$  (resp.  $\Gamma_N$ ) contains edges lying on the boundary where Dirichlet (resp. Neumann) boundary conditions are imposed and in  $\Gamma_I$  there are all internal edges, which we call interfaces, as they frequently correspond to the physical interfaces between different semiconductor materials. We also define

$$\Gamma_{DI} := \Gamma_D \cup \Gamma_I, \quad \Gamma_i := \{e \in \Gamma : e \subset \partial\Omega_i\}. \quad (9)$$

Let  $e \in \Gamma$ . Then two cases are possible. Either  $e \in \Gamma_D \cup \Gamma_N$ , so there is an unique  $\Omega_i$  such that  $e$  is an edge of  $\Omega_i$ , or  $e \in \Gamma_I$  and there are exactly two sets  $\Omega_i, \Omega_j \in \mathcal{E}$  such that  $e$  is their common edge. We will often refer to these two cases. Also we define  $\text{nb}(\Omega_i) := \{\Omega_k \in \mathcal{E} : \Gamma_i \cap \Gamma_k \neq \emptyset\}$ .

For  $s > 1/2$  we define operators  $[\cdot] := [\cdot]_e : H^s(\mathcal{E}) \rightarrow L_2(e)$ ,  $\{\cdot\} := \{\cdot\}_e : H^s(\mathcal{E}) \rightarrow L_2(e)$  as

$$\begin{aligned} [u] &:= \begin{cases} u_i - u_j & \text{if } e \in \Gamma_I, e = \partial\Omega_i \cap \partial\Omega_j, \\ u_i & \text{if } e \in \Gamma_D \cup \Gamma_N, e = \partial\Omega_i \cap \partial\Omega, \end{cases} \\ \{u\} &:= \begin{cases} \frac{1}{2}(u_i + u_j) & \text{if } e \in \Gamma_I, e = \partial\Omega_i \cap \partial\Omega_j, \\ u_i & \text{if } e \in \Gamma_D \cup \Gamma_N, e = \partial\Omega_i \cap \partial\Omega. \end{cases} \end{aligned} \quad (10)$$

For convenience, we will also use this notion for triangulation parameters, i.e.

$$\{h^{-s}\} := \left\{ \frac{1}{h^s} \right\} := \begin{cases} \frac{1}{2} \left( \frac{1}{h_i^s} + \frac{1}{h_j^s} \right) & \text{if } e = \partial\Omega_i \cap \partial\Omega_j, \\ \frac{1}{h_i^s} & \text{if } e = \partial\Omega_i \cap \partial\Omega. \end{cases} \quad (11)$$

For further analysis, we introduce so-called “broken” norm  $\|\cdot\|_{h,\sigma_r}$  in  $X_h(\Omega)$  with

$$\|u_h\|_{h,\sigma_r}^2 := \sum_{i=1}^N \int_{\Omega_i} \varepsilon (\nabla u_{h,i})^2 dx + \sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u_h]^2 ds. \quad (12)$$

Here  $\eta_{r,e}$  is a penalty coefficient for  $e$ . It depends on the triangulation parameters and penalty parameters  $\sigma_e > 0$ :

$$\eta_{r,e} := 2\sigma_e \{h^{-r}\} = \begin{cases} 2\sigma_e h_i^{-r} & e \in \Gamma_D, e \subset \Omega_i, \\ \sigma_e (h_i^{-r} + h_j^{-r}) & e \in \Gamma_I, e \subset \Omega_i \cap \Omega_j. \end{cases} \quad (13)$$

Depending on the method, we will use  $r = 1$  or  $r = 2$ .

To simplify the analysis, we assume that  $0 < \sigma_0 \leq \sigma_e$  for all  $e \in \Gamma_{DI}$ . Also we assume that  $0 < h_i < h_0 \leq 1$  for all  $i \in \{1, \dots, N\}$ . The choice of  $\sigma_0$  and  $h_0$  will be discussed later in lemmata 3.2, 3.3 and 7.2.

**Lemma 3.1.** *For any  $u_h \in X_h(\Omega)$ ,  $\Omega_i \in \mathcal{E}$  and  $e \in \Gamma_i$ , the following estimates hold*

$$\|u_{h,i}\|_{L_2(e)} \leq Ch_i^{-1/2} \|u_h\|_{L_2(\Omega_i)}, \quad (14)$$

$$\|\nabla u_{h,i} \cdot \nu\|_{L_2(e)} \leq Ch_i^{-1/2} |u_h|_{H^1(\Omega_i)}. \quad (15)$$

Constant  $C$  does not depend on  $h_i$ .

*Proof.* These estimates are a consequence of the trace theorem applied to each edge of fine elements in  $\Omega_i$  coincident with  $e$  followed by a scaling argument.  $\square$

## 3.2 Composite Discontinuous Galerkin variants

We propose two variants of the Composite Discontinuous Galerkin discretization. First one is based on Weakly Over-Penalized Symmetric Interior Penalty (WOPSIP) method (cf. [3] or [2]). Second formulation is derived from Symmetric Interior Penalty Galerkin (SIPG) method (cf. [13] or [12]). In each case we use the composite formulation (cf. [7]), i.e. inside every  $\Omega_i$  we use the Finite Element Method on the triangulation  $\mathcal{T}_{h_i}$ , while on boundaries  $e \in \Gamma_{DI}$  we use the respective variant of the Discontinuous Galerkin Method. Thus we call these methods Composite Weakly Over-Penalized Symmetric Interior Penalty (CWOPSIP) method and Composite Symmetric Interior Penalty Galerkin (CSIPG) method, respectively.

### 3.2.1 Composite Weakly Over-Penalized Symmetric Interior Penalty (CWOPSIP)

This discretization has a simpler formulation from the two methods we introduce. The discrete problem is defined as follows. Find  $u_h^* \in X_h(\Omega)$  such that

$$a_{2,h}(u_h^*, \varphi_h) + b(u_h^*, \varphi_h) = f_{2,h}(\varphi_h), \quad \forall \varphi_h \in X_h(\Omega), \quad (16)$$

where

$$\begin{aligned} a_{2,h}(u_h, \varphi_h) &= \sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u_{h,i}^* \cdot \nabla \varphi_{h,i} dx + \sum_{e \in \Gamma_{DI}} \eta_{2,e} \int_e [u_h] \cdot [\varphi_h] ds, \\ f_{2,h}(\varphi_h) &= \int_{\Omega} k_1 \varphi_h dx + \sum_{e \in \Gamma_D} \eta_{2,e} \int_e [\hat{u}] \cdot [\varphi_h] ds, \end{aligned} \quad (17)$$

and  $b$  is defined as in (3).

### 3.2.2 Composite Symmetric Interior Penalty Galerkin (CSIPG)

This problem is defined as follows. Find  $u_h^* \in X_h(\Omega)$  such that

$$a_{1,h}(u_h^*, \varphi_h) + b(u_h^*, \varphi_h) = f_{1,h}(u_h^*, \varphi_h), \quad \forall \varphi_h \in X_h(\Omega), \quad (18)$$

where

$$\begin{aligned} a_{1,h}(u_h, \varphi_h) &= \sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u_{h,i}^* \cdot \nabla \varphi_{h,i} dx - \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla u_h \cdot \nu\} [\varphi_h] ds \\ &\quad - \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla \varphi_h \cdot \nu\} [u_h] ds + \sum_{e \in \Gamma_{DI}} \eta_{1,e} \int_e [u_h] \cdot [\varphi_h] ds, \\ f_{1,h}(\varphi_h) &= \int_{\Omega} k_1 \varphi_h dx - \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla \varphi_h \cdot \nu\} [\hat{u}] ds \\ &\quad + \sum_{e \in \Gamma_D} \eta_{1,e} \int_e [\hat{u}] [\varphi_h] ds, \end{aligned} \quad (19)$$

and  $b$  is defined as in (3).

**Lemma 3.2.** *For any  $\alpha \in (0, 1)$  there exist  $\sigma_0 > 0$ , such that for every  $\sigma_e \geq \sigma_0$  and  $u_h \in X_h(\Omega)$*

$$2 \sum_{e \in \Gamma_{DI}} \left| \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \right| \leq \alpha \|u_h\|_{h, \sigma_1}^2, \quad (20)$$

*Proof.* Let us take any  $e = \Omega_j \cap \Omega_k$ . By the Schwarz inequality and a triangle inequality

$$\begin{aligned} \left| 2 \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \right| &\leq \left( \|\varepsilon|_{\Omega_j} \nabla u_{h,j} \cdot \nu\|_{L_2(e)} \right. \\ &\quad \left. + \|\varepsilon|_{\Omega_k} \nabla u_{h,k} \cdot \nu\|_{L_2(e)} \right) \| [u_h] \|_{L_2(e)} \end{aligned} \quad (21)$$

Taking  $\Omega_i \in \{\Omega_j, \Omega_k\}$ , we use the Cauchy's  $\epsilon$ -inequality and lemma 3.1

$$\begin{aligned} &\|\varepsilon|_{\Omega_i} \nabla u_{h,i} \cdot \nu\|_{L_2(e)} \| [u_h] \|_{L_2(e)} \\ &\leq \epsilon \varepsilon_M h_i \|\nabla u_{h,E}\|_{L_2(e)}^2 + \frac{1}{4\epsilon} \varepsilon_M \frac{1}{h_i} \| [u_h] \|_{L_2(e)}^2 \\ &\leq \epsilon \varepsilon_M \|\nabla u_h\|_{L_2(\Omega_i)}^2 + \frac{1}{4\epsilon} \varepsilon_M \frac{1}{h_i} \| [u_h] \|_{L_2(e)}^2. \end{aligned} \quad (22)$$

Therefore we get

$$\left| 2 \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \right| \leq 2\varepsilon_M \epsilon \|\nabla u_h\|_{L_2(\Omega_j)}^2 + \varepsilon_M \frac{1}{2\epsilon} \{h^{-1}\} \| [u_h] \|_{L_2(e)}^2. \quad (23)$$

On the other hand, if  $e \in \Gamma_D$  then  $e \in \partial\Omega_i \cap \partial\Omega$  and by similar arguments we have

$$\left| 2 \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \right| \leq 2\varepsilon_M \epsilon \|\nabla u_h\|_{L_2(\Omega_i)}^2 + \varepsilon_M \frac{1}{\epsilon} \{h^{-1}\} \| [u_h] \|_{L_2(e)}^2 \quad (24)$$

Summing these results up we get

$$\begin{aligned} 2 \sum_{e \in \Gamma_{DI}} \left| \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \right| &\leq \epsilon C \sum_{i=1}^N \|\nabla u_h\|_{L_2(\Omega_i)}^2 \\ &+ \frac{C}{\epsilon} \sum_{e \in \Gamma_{DI}} \{h^{-1}\} \| [u_h] \|_{L_2(e)}^2, \end{aligned} \quad (25)$$

where  $C$  depends on  $\varepsilon_M$  and the maximal number of edges of elements in coarse grid  $\mathcal{E}$ . Finally for any  $\alpha \in (0, 1)$  if we take  $\epsilon := \alpha \varepsilon_m / C$  and then taking  $\sigma_0(\alpha) := C^2 / \alpha^2 \varepsilon_m$  we have

$$\begin{aligned} 2 \sum_{e \in \Gamma_{DI}} \left| \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \right| &\leq \alpha \sum_{i=1}^N \varepsilon_m \|\nabla u_h\|_{L_2(\Omega_i)}^2 \\ &+ \alpha \sum_{e \in \Gamma_{DI}} \frac{C^2}{\varepsilon_m \alpha^2} \{h^{-1}\} \| [u_h] \|_{L_2(e)}^2 \\ &\leq \alpha \|u_h\|_{h, \sigma_1}^2. \end{aligned} \quad (26)$$

□

**Lemma 3.3.** *There exist  $\sigma_0 > 0$  and  $c > 0$ , such that for every  $\sigma_e \geq \sigma_0$  and  $u_h \in X_h(\Omega)$*

$$c \|u_h\|_h^2 \leq a_{h,1}(u_h, u_h). \quad (27)$$

*Proof.* Due to definition of broken norm, we have

$$a_{h,1}(u_h, u_h) = \|u_h\|_{h, \sigma_1}^2 - 2 \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds. \quad (28)$$

Let us take  $\alpha := 1/2$  in lemma 3.2. Then we have

$$\begin{aligned} a_{h,1}(u_h, u_h) &= \|u_h\|_{h, \sigma_1}^2 - 2 \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla u_h \cdot \nu\} [u_h] ds \\ &\geq \|u_h\|_{h, \sigma_1}^2 - \frac{1}{2} \|u_h\|_{h, \sigma_1}^2 \geq \frac{1}{2} \|u_h\|_{h, \sigma_1}^2. \end{aligned} \quad (29)$$

□



## 4 Existence

Let us denote the nodal base of  $X_h$  by  $\{\varphi_{(j)}\}_{i=1}^J$ . Then we define  $P : X_h(\Omega) \rightarrow X_h^*(\Omega)$  as

$$P(u_h)\varphi_h := a_{r,h}(u_h, \varphi_h) + b(u_h, \varphi_h) - f_{r,h}(\varphi_h). \quad (30)$$

We would like to use the following theorem [10]:

**Theorem 4.1.** (Brouwer) *Let  $P : X \rightarrow X^*$  be a continuous function on a finite-dimensional normed real vector space  $X$ , such that for suitable  $\rho > 0$  we have*

$$P(x)x \geq 0 \quad \forall \|x\| \geq \rho, \quad (31)$$

*Then there exists  $x \in X$ ,  $\|x\| \leq \rho$  such that*

$$P(x) = 0. \quad (32)$$

### 4.1 Case of CWOPSIP

By definition (30)

$$P(u_h)u_h := a_{2,h}(u_h, u_h) + b(u_h, u_h) - f_{2,h}(u_h), \quad (33)$$

Then we have that

$$a_{2,h}(u_h, u_h) = \|u_h\|_{h,\sigma_2}^2. \quad (34)$$

On the other hand, by Schwarz inequality

$$-f_{2,h}(u_h) \geq -c(\hat{u}, k_1)\|u_h\|_{L_2(\Omega)} \geq -c(\hat{u}, k_1, h)\|u_h\|_{h,\sigma_2} \quad (35)$$

Then let  $C := \max\{\|\hat{v}\|_{L_\infty(\Omega)}, \|\hat{w}\|_{L_\infty(\Omega)}\}$ . Then we may decompose  $\Phi_2$  as

$$\begin{aligned} b(u_h, u_h) &= \int_{\Omega} \left( e^{u_h - \hat{v}} - e^{\hat{w} - u_h} \right) u_h dx \\ &= \int_{\Omega} \left( e^{u_h - \hat{v}} - e^{\hat{w} - u_h} \right) u_h \chi_{\{x \in \Omega : |u_h(x)| > C\}} dx \\ &\quad + \int_{\Omega} \left( e^{u_h - \hat{v}} - e^{\hat{w} - u_h} \right) u_h \chi_{\{x \in \Omega : |u_h(x)| \leq C\}} dx \end{aligned} \quad (36)$$

The first integral is non-negative, and the latter we can estimate from below

$$\int_{\Omega} \left( e^{u_h(x) - \hat{v}(x)} - e^{\hat{w}(x) - u_h(x)} \right) u_h(x) \chi_{\{x \in \Omega : |u_h(x)| \leq C\}}(x) dx \geq -|\Omega| 2e^{2C} C. \quad (37)$$

In conclusion, we may use these estimations to obtain

$$P(u_h)u_h \geq \|u_h\|_{h,\sigma_2}^2 - c_1 \|u_h\|_{h,\sigma_2} - c_2 \quad (38)$$

Note that constants  $c_i$  in this inequality depend on  $h$ .

It is therefore clear that for  $\|u_h\|_{h,\sigma_2}$  large enough, we have that  $P(u_h)u_h \geq 0$ . Then by theorem 4.1 we have that there exists some  $u_h^*$ , such that  $P(u_h^*) = 0$ .

## 4.2 Case of CSIPG

We proceed analogously to the CWOPSIP case. For  $b(u_h, u_h)$  the argumentation is exactly the same. Then  $f_{1,h}(u_h)$  has one additional term, which may be estimated using lemma 3.1 and the trace inequality

$$\begin{aligned} \left| \sum_{e \in \Gamma_D} \int_e \{\varepsilon \nabla u_h \cdot \nu\} [\hat{u}] ds \right| &\leq \varepsilon_M \sum_{e \in \Gamma_D} \|\nabla u_h\|_{L_2(e)} \|\hat{u}\|_{L_2(e)} \\ &\leq c\varepsilon_M \sum_{i=1}^N h_i^{-1/2} \|\nabla u_h\|_{L_2(\Omega_i)} \|[\hat{u}]\|_{L_2(\Omega_i)} \\ &\leq C \|u_h\|_{h, \sigma_1} \|\hat{u}\|_{L_2(\Omega)}, \end{aligned} \quad (39)$$

where  $C$  depends on  $\varepsilon_M$  and  $h$ . Then estimating  $a_{1,h}(u_h, u_h)$  by lemma (3.3), we have

$$P(u_h)u_h \geq c_1 \|u_h\|_{h, \sigma_1}^2 - c_2 \|u_h\|_{h, \sigma_1} - c_3 \quad (40)$$

The existence of  $u_h^*$  is now proven.

## 5 Uniqueness

The uniqueness can be shown by contradiction for both cases. Here we present the CWOPSIP case, for CSIPG the proof is similar.

Assume that there exist two solutions  $u_h, v_h \in X_h(\Omega)$  of equation (18). Then by taking  $\varphi_h := u_h - v_h$  in (18) and subtracting these equations for  $u_h$  and  $v_h$  we obtain

$$\begin{aligned} a_{2,h}(u_h - v_h, u_h - v_h) &= \sum_{i=1}^N \int_{\Omega_i} e^{-\hat{v}} (e^{v_h} - e^{u_h}) (u_h - v_h) dx \\ &\quad + \sum_{i=1}^N \int_{\Omega_i} e^{\hat{w}} (e^{-u_h} - e^{-v_h}) (u_h - v_h) dx. \end{aligned} \quad (41)$$

By monotonicity of the exponential function, the right hand side is nonpositive. On the other hand

$$0 < \|u_h - v_h\|_{h, \sigma_2}^2 = a_{2,h}(u_h - v_h, u_h - v_h). \quad (42)$$

Thus  $0 < \|u_h - v_h\|_{h, \sigma_2}^2 \leq 0$  since  $u_h \neq v_h$ , and we have a contradiction.

## 6 Interpolation operator

First let us take any  $\Omega_i \in \mathcal{E}$  and let us define interpolation operator  $\Pi_{h_i} : H^2(\Omega_i) \rightarrow X_{h_i} \subset C^0(\overline{\Omega_i})$  as follows

$$\forall x \in \mathcal{N}_{h_i} \quad \Pi_{h_i} u(x) = u(x). \quad (43)$$

Note that for  $s > 0$  we have  $H^{1+s}(\Omega) \subset C^0(\Omega)$  (see [13]), so this definition is well-posed. Then we define  $\Pi_h : H^2(\mathcal{E}) \rightarrow X_h$  by

$$\forall \Omega_i \in \mathcal{E} \quad \left( \Pi_h u \right)_i := \Pi_{h_i} u_i. \quad (44)$$

On any  $\Omega_i$ , we can use standard interpolation estimate for FEM [5]

$$\|u_i - \Pi_{h_i} u_i\|_{L_2(\Omega_i)} + h_i \|u_i - \Pi_{h_i} u_i\|_{H^1(\Omega_i)} \leq Ch_i^2 |u_i|_{H^2(\Omega_i)}. \quad (45)$$

We remind that coarse grid  $\mathcal{E}$  is independent of  $h$ . Thus

$$\|u - \Pi_h u\|_{H^1(\Omega)}^2 = \sum_{i=1}^N \|u_i - \Pi_{h_i} u_i\|_{H^1(\Omega_i)}^2 \leq C \sum_{\Omega_i \in \mathcal{E}} h_i^2 |u_i|_{H^2(\Omega_i)}^2. \quad (46)$$

Let further  $u_I := \Pi_h u$ .

**Lemma 6.1.** *Let  $u \in H^2(\mathcal{E})$ ,  $u_I := \Pi_h u$ . For any  $\Omega_i \in \mathcal{E}$  and for any  $e \in \Gamma_i$*

$$\|u_i - u_{I,i}\|_{L_2(e)} \leq Ch_i^{3/2} |u_i|_{H^2(\Omega_i)}, \quad (47)$$

$$|u_i - u_{I,i}|_{H^1(e)} \leq Ch_i^{1/2} |u_i|_{H^2(\Omega_i)}. \quad (48)$$

*Proof.* For fixed  $e$  and  $\Omega_i$ , we have  $\|u_i - u_{I,i}\|_{L_2(e)}^2 = \sum_{\tau \in \mathcal{T}_{h_i,e}} \|u_i - u_{I,i}\|_{L_2(e \cap \tau)}^2$ . Note that on a single triangulation element  $\tau$  we have that  $u_i - u_{I,i} \in H^2(\tau)$ , so using the trace inequality (see [13]) for  $H^2(\tau)$  functions we have

$$\|u_i - u_{I,i}\|_{L_2(e \cap \tau)} = C \left( h_i^{-1/2} \|u_i - u_{I,i}\|_{L_2(\tau)} + h_i^{1/2} |u_i - u_{I,i}|_{H^1(\tau)} \right). \quad (49)$$

Then by (45) it follows that

$$\begin{aligned} \|u_i - u_{I,i}\|_{L_2(e)} &\leq Ch_i^{-1/2} \left( \|u_i - u_{I,i}\|_{L_2(\Omega_i)} + h_i \|u_i - u_{I,i}\|_{H^1(\Omega_i)} \right) \\ &\leq Ch_i^{3/2} |u_i|_{H^2(\Omega_i)}. \end{aligned} \quad (50)$$

Proof of the latter estimate is analogous.  $\square$

Let us take any  $e \in \Gamma_{DI}$ . For  $e \in \Gamma_I$  we assume that  $e = \Omega_j \cap \Omega_k$  for some  $\Omega_j, \Omega_k \in \mathcal{E}$  and by a triangle inequality we have

$$\int_e [u - u_I]^2 ds = \|[u - u_I]\|_{L_2(e)}^2 \leq 2 \sum_{i \in \{j,k\}} \|u_i - u_{I,i}\|_{L_2(e)}^2, \quad (51)$$

while for  $e \in \Gamma_D$  we have  $e \in \Gamma_i$  for some  $\Omega_i \in \mathcal{E}$  and simply

$$\int_e [u - u_I]^2 ds = \int_e (u_i - u_{I,i})^2 ds = \|u_i - u_{I,i}\|_{L_2(e)}^2 \quad (52)$$

Therefore it is sufficient to estimate  $\|u_i - u_{I,i}\|_{L_2(e)}^2$ , for any  $e \in \Gamma_{DI}$ ,  $e \subset \partial\Omega_i$ , as in lemma 6.1

To proceed further, we need to take into account  $\eta_{r,e}$ . Therefore we will distinguish few cases.

**CSIPG** In this case  $\eta_{1,e} = \sigma_e \{h^{-1}\}$ . Let  $e \in \Gamma_D$ . By (50) we have

$$\begin{aligned} \eta_{1,e} \int_e \left(u_i - u_{I,i}\right)^2 ds &= \sigma_e h_i^{-1} \|u_i - u_{I,i}\|_{L_2(e)}^2 \\ &\leq C \sigma_e h_i^{-1} h_i^3 |u_i|_{H^2(\Omega_i)}^2 = C \sigma_e h_i^2 |u_i|_{H^2(\Omega_i)}^2. \end{aligned} \quad (53)$$

On the other hand, if  $e \in \Gamma_I$  then

$$\begin{aligned} \eta_{1,e} \int_e \left(u_j - u_{I,j}\right)^2 ds &= 0.5 \sigma_e (h_j^{-1} + h_k^{-1}) \|u_j - u_{I,j}\|_{L_2(e)}^2 \\ &\leq C \sigma_e \left(h_j^2 + \frac{h_j^3}{h_k}\right) |u_j|_{H^2(\Omega_j)}^2. \end{aligned} \quad (54)$$

Then if we sum up over  $e \in \Gamma_{DI}$

$$\sum_{e \in \Gamma_{DI}} \eta_{1,e} \int_e \left([u - u_I]\right)^2 ds \leq \sum_{i=1}^N C \left(h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k}\right) |u_i|_{H^2(\Omega_i)}^2. \quad (55)$$

Thus taking into account this estimate and (46)

$$\|u - u_I\|_{h, \sigma_1}^2 \leq C \sum_{i=1}^N \left(h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k}\right) |u_i|_{H^2(\Omega_i)}^2. \quad (56)$$

If we increase destiny poportionally, i.e.  $h_i := c_i h$ , the result can be improved to

$$\|u - u_I\|_{h, \sigma_1}^2 \leq C h^2 \sum_{i=1}^N |u_i|_{H^2(\Omega_i)}^2. \quad (57)$$

**CWOPSIP - general case** Here we have  $\eta_{2,e} = \sigma_e \{h^{-2}\}$ . Proceeding as in previous case, we get for  $e \in \Gamma_D$

$$\eta_{2,e} \int_e \left(u_i - u_{I,i}\right)^2 ds \leq C \sigma_e h_i |u_i|_{H^2(\Omega_i)}^2, \quad (58)$$

and for  $e \in \Gamma_I$

$$\eta_{2,e} \int_e \left(u_j - u_{I,j}\right)^2 ds \leq C \sigma_e \left(h_j + \frac{h_j^3}{h_k^2}\right) |u_j|_{H^2(\Omega_j)}^2. \quad (59)$$

So taking into account (46)

$$\|u - u_I\|_{h, \sigma_2}^2 \leq C \sum_{i=1}^N \left(h_i + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k^2}\right) |u_i|_{H^2(\Omega_i)}^2. \quad (60)$$

Again assuming  $h_i := c_i h$  we obtain

$$\|u - u_I\|_{h, \sigma_2}^2 \leq C h \sum_{i=1}^N |u_i|_{H^2(\Omega_i)}^2. \quad (61)$$

We must note that this estimate is generally worse than for CSIPG method.

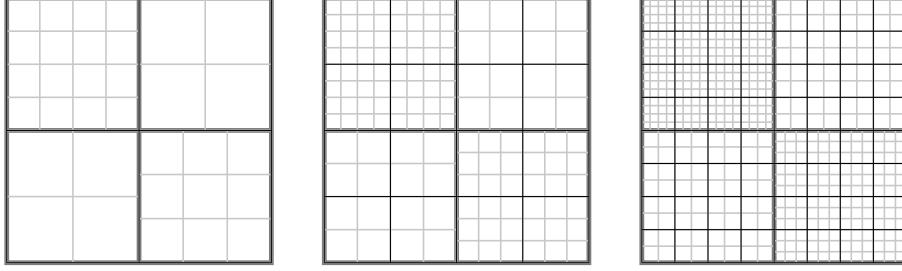


Figure 3: An example of a mesh of  $V_h(\Omega)$  space. Fine grid (mesh of  $X_h(\Omega)$ ) lines are in light gray, coarse grid ( $\mathcal{E}$ ) lines are in dark gray and the black color is used for mesh of  $V_h(\Omega)$ . Three levels of refinement are presented.

**CWOPSIP - special case** Imposing some restrictions on the mesh and boundary conditions, we can improve the interpolation error estimate. Assume that for every space  $X_h(\Omega)$  there is some continuous FEM space  $V_h(\Omega) \subset \mathcal{C}(\Omega)$ , such that  $V_h(\Omega) \subset X_h(\Omega)$ . An example is presented on figure 3.

Under such assumptions, we define the projection operator  $\Pi_h : H^2(\mathcal{E}) \rightarrow V_h(\Omega) \subset X_h(\Omega)$  to satisfy  $\Pi_h u(x) = u(x)$  for all nodal points of discretization  $V_h$ . Then  $\Pi_h u = u_I$  is continuous and  $[u_I]_e \equiv 0$  for every  $e \in \Gamma_I$ . Also we have  $[u]_e \equiv 0$  in this case as  $u \in H^1(\Omega)$  and

$$\sum_{e \in \Gamma_{DI}} \eta_{2,e} \int_e ([u - u_I])^2 ds = \sum_{e \in \Gamma_D} \eta_{2,e} \int_e ([\hat{u} - \hat{u}_I])^2 ds. \quad (62)$$

This integral do not vanish in general. However we additionally assume that  $\hat{u}|_e$  is constant on any  $e \in \Gamma_D$ . In this paper, we deal with electrostatic potential, which do not vary within a given contact of a device. Thus by choosing the division in such a manner that boundary edges  $e \in \Gamma_D$  do not cross the contact boundaries, we can satisfy this assumption. Then  $\hat{u}|_e = u_I|_e$  and simply by (46)

$$\begin{aligned} \|u - u_I\|_{h,\sigma_2}^2 &= \sum_{i=1}^N \int_{\Omega_i} \varepsilon_i \nabla(u - u_I)^2 dx \leq C \sum_{i=1}^N h_i^2 |u_i|_{H^2(\Omega_i)}^2 \\ &\leq Ch^2 \sum_{i=1}^N |u_i|_{H^2(\Omega_i)}. \end{aligned} \quad (63)$$

## 7 Error estimates

### 7.1 Definitions and additional assumptions

We start with auxiliary lemma.

**Lemma 7.1.** *Let  $u \in H^s(\mathcal{E})$ ,  $s \geq 1$ . Then*

$$\|u\|_{L_2(\Omega)}^2 \leq C \left[ \sum_{i=1}^N \int_{\Omega_i} (\nabla u)^2 dx + \sum_{e \in \Gamma_I} |e|^{-1} \int_e [u]^2 ds + \sum_{e \in \Gamma_D} \int_e u^2 ds \right]. \quad (64)$$

Proof of lemma 7.1 may be found in [4].

Then we would like to have an analogue of a Poincare inequality for the  $H^1(\mathcal{E})$  space.

**Lemma 7.2.** *Let  $u \in H^s(\mathcal{E})$ ,  $s \geq 1$ ,  $r \in \{1, 2\}$ . Then there exists some  $h_0 > 0$ , such that  $\|u\|_{L_2(\Omega)} \leq c\|u\|_{h, \sigma_r}$  for  $0 < h \leq h_0$ , where  $c$  is independent of  $h$ .*

*Proof.* By definition of the broken norm (12), we have

$$\|u\|_{h, \sigma_r}^2 := \sum_{i=1}^N \int_{\Omega_i} \varepsilon_i (\nabla u)^2 dx + \sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u]^2 ds. \quad (65)$$

Note that  $|e|$  do not depend on  $h$  and  $\eta_{r,e} \rightarrow \infty$  as  $h \rightarrow 0$ . Thus we can find  $h_0 > 0$ , such that  $\eta_{r,e} \geq |e|^{-1}$  and  $\eta_{r,e} \geq 1$  for any  $0 < h < h_0$  and then by lemma 7.1

$$\begin{aligned} \|u\|_{L_2(\Omega)}^2 &\leq C \left[ \sum_{i=1}^N \int_{\Omega_i} (\nabla u)^2 dx + \sum_{e \in \Gamma_I} |e|^{-1} \| [u] \|_{L_2(e)}^2 + \sum_{e \in \Gamma_D} \|u\|_{L_2(e)}^2 \right] \\ &\leq C \left[ \varepsilon_m^{-1} \sum_{i=1}^N \int_{\Omega_i} \varepsilon_i (\nabla u)^2 dx + \sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u]^2 ds \right] \\ &\leq C_1 \|u\|_{h, \sigma_r}^2. \end{aligned} \quad (66)$$

□

To prove error estimates of the proposed discretizaion, we would like to introduce the following assumptions:

$$u^* \in H^1(\Omega) \cap H^2(\mathcal{E}), \quad \varepsilon \in \{v \in L_\infty(\Omega) : \forall i \in \{1, \dots, n\} v|_{\Omega_i} \in \mathcal{C}^1(\overline{\Omega})\}. \quad (67)$$

## 7.2 Consistency

We start with an abstract result. Let  $u \in H^1(\Omega)$ ,  $f \in L_2(\Omega)$ . We pose two problems. The first is the following: find  $u \in H^1(\Omega)$  such that

$$\begin{aligned} \int_{\Omega} \varepsilon \nabla u \cdot \nabla \varphi dx &= \int_{\Omega} f \varphi dx \quad \forall \varphi \in H_{0, \Sigma_D}^1(\Omega), \\ u &= \hat{u} \quad \text{on } \Sigma_D. \end{aligned} \quad (68)$$

Second problem is posed in “broken” Sobolev space: find  $u \in H^1(\mathcal{E})$ , such that  $\forall \varphi \in H^1(\mathcal{E}) \cap H^2(\mathcal{T}_h)$

$$\begin{aligned} &\sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi dx - \sum_{e \in \Gamma_{DI}} \int_e \{ \varepsilon \nabla u \cdot \nu \} [\varphi] ds \\ &\quad + \xi \sum_{e \in \Gamma_{DI}} \int_e \{ \varepsilon \nabla \varphi \cdot \nu \} [u] ds + \sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u][\varphi] ds \\ &= \sum_{i=1}^N \int_{\Omega_i} f \varphi dx + \xi \sum_{e \in \Gamma_D} \int_e \{ \varepsilon \nabla \varphi \cdot \nu \} [\hat{u}] ds + \sum_{e \in \Gamma_D} \eta_{r,e} \int_e [\hat{u}][\varphi] ds. \end{aligned} \quad (69)$$

We consider  $\xi \in \{0, -1\}$ , which will be used in CWOPSIP and CSIPG, respectively.

We would like to prove the following result.

**Theorem 7.3.** *Assume that the solution  $u$  of problem (68) belongs to  $H^1(\Omega) \cap H^2(\mathcal{E})$  and  $\varepsilon \nabla u \in H^1(\mathcal{E})$ . Then  $u$  satisfies (69). Conversely, if  $u \in H^2(\mathcal{E}) \cap H^1(\Omega)$  is a solution of (69) and  $\varepsilon \nabla u \in H^1(\mathcal{E})$ , then it is also a solution of (68).*

The proof is based on the standard approach in Discontinuous Galerkin Method [13].

**Lemma 7.4.** *Let  $u \in H^1(\Omega) \cap H^2(\mathcal{E})$ ,  $\varepsilon \nabla u \in (H^1(\mathcal{E}))^2$ ,  $0 < \varepsilon_m \leq \varepsilon \leq \varepsilon_M$  and  $f \in L_2(\Omega)$ . The following statements are equivalent:*

- $u$  satisfy:

$$\int_{\Omega} \varepsilon \nabla u \cdot \nabla \varphi = \int_{\Omega} f \varphi, \quad \forall \varphi \in H_{0,\Sigma_D}^1(\Omega) \quad (70)$$

- $u$  satisfy:

$$\begin{aligned} - \sum_{i=1}^N \int_{\Omega_i} \nabla(\varepsilon_i \nabla u_i) \varphi_i &= \int_{\Omega} f \varphi, \quad \forall \varphi \in L_2(\Omega) \\ [\varepsilon \nabla u \cdot \nu] &= 0 \quad \text{on } \Sigma_I, \\ \nabla u \cdot \nu &= 0 \quad \text{on } \Sigma_N. \end{aligned} \quad (71)$$

*Proof.* (71)  $\Rightarrow$  (70) follows simply from the Green formula. To prove (70)  $\Rightarrow$  (71), take any  $\varphi \in C_0^\infty(\Omega)$ . Since  $C_0^\infty(\Omega) \subset H_{0,\Sigma_D}^1(\Omega)$ , then by (70) we have

$$\int_{\Omega} \varepsilon \nabla u \cdot \nabla \varphi = \int_{\Omega} f \varphi. \quad (72)$$

By the Green formula

$$\begin{aligned} \int_{\Omega} f \varphi dx &= \sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi dx \\ &= - \sum_{i=1}^N \int_{\Omega_i} \nabla(\varepsilon \nabla u) \varphi dx + \sum_{e \in \Gamma} \int_e [\varepsilon \nabla u \cdot \nu] \varphi ds \end{aligned} \quad (73)$$

Since  $\varphi$  is zero on  $\partial\Omega$ , we may rewrite last sum

$$\int_{\Omega} f \varphi = - \sum_{i=1}^N \int_{\Omega_i} \nabla(\varepsilon \nabla u) \varphi dx + \sum_{e \in \Gamma_I} \int_e [\varepsilon \nabla u \cdot \nu] \varphi ds. \quad (74)$$

Then, as a simple consequence of the Green formula, for every  $e \in \Gamma_I \cup \Gamma_N$  we have  $[\varepsilon \nabla u \cdot \nu] = 0$ , so

$$- \sum_{i=1}^N \int_{\Omega_i} \nabla(\varepsilon \nabla u) \varphi dx = \int_{\Omega} f \varphi. \quad (75)$$

This statement is true for  $\varphi \in C_0^\infty(\Omega)$ . Since  $\nabla(\varepsilon \nabla u) \in L_2(\Omega)$  it is also true for any  $\varphi \in L_2(\Omega)$  and first statement of (71) is shown.  $\square$

*Proof.* (Theorem 7.3)

First we prove (68)  $\Rightarrow$  (69). Assume that  $u$  is a solution of (68) and that it belongs to  $H^1(\Omega) \cap H^2(\mathcal{E})$ . We have by definition

$$\int_{\Omega} \varepsilon \nabla u \cdot \nabla \phi \, dx = \int_{\Omega} f \phi \, dx \quad \forall \phi \in H_{0,\Sigma_D}^1(\Omega). \quad (76)$$

We use lemma 7.4 and we obtain that for any  $\phi \in L_2(\Omega)$

$$-\int_{\Omega} \nabla \cdot (\varepsilon \nabla u) \phi \, dx = \int_{\Omega} f \phi \, dx. \quad (77)$$

Let us take any  $\varphi \in H^1(\mathcal{E}) \cap H^2(\mathcal{T}_h)$  and substitute  $\phi := \varphi$ . We may split integrals to

$$-\sum_{i=1}^N \int_{\Omega_i} \nabla \cdot (\varepsilon \nabla u) \varphi \, dx = \sum_{i=1}^N \int_{\Omega_i} f \varphi \, dx. \quad (78)$$

By the Green theorem, we have

$$\int_{\Omega_i} \nabla \cdot (\varepsilon \nabla u) \varphi \, dx = \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi \, dx - \int_{\partial \Omega_i} \varepsilon \nabla u \cdot \nu \varphi \, dx. \quad (79)$$

Summing up these results in  $\Omega_i$ , we get

$$\sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi \, dx - \sum_{i=1}^N \int_{\partial \Omega_i} \varepsilon \nabla u \cdot \nu \varphi \, dx = \sum_{i=1}^N \int_{\Omega_i} f \varphi \, dx. \quad (80)$$

By lemma 7.4, we have that  $[\varepsilon \nabla u] = 0$  on  $\Sigma_I$ , thus  $\{\varepsilon \nabla u \cdot \nu\} = \varepsilon \nabla u \cdot \nu$  on any  $\partial \Omega_i$  and we have

$$\sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi \, dx - \sum_{e \in \Gamma} \int_e \{\varepsilon \nabla u \cdot \nu\} [\varphi] \, dx = \sum_{i=1}^N \int_{\Omega_i} f \varphi \, dx. \quad (81)$$

By homogeneous Neumann boundary condition (lemma 7.4) on  $e \in \Gamma_N$  we have  $\{\varepsilon \nabla u \cdot \nu\} = 0$  and

$$\sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi \, dx - \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla u \cdot \nu\} [\varphi] \, dx = \sum_{i=1}^N \int_{\Omega_i} f \varphi \, dx. \quad (82)$$

Since  $u \in H^1(\Omega)$ , then  $[u] = 0$  for any  $e \in \Gamma_I$  and by assumption on  $e \in \Gamma_D$  we have  $u = \hat{u}$  so we have for any  $\varphi \in H^1(\mathcal{E})(\Omega)$

$$\begin{aligned} \sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u] [\varphi] \, ds + \xi \sum_{e \in \Gamma_{DI}} \int_e \{\varepsilon \nabla \varphi \cdot \nu\} [u] \, ds \\ = \sum_{e \in \Gamma_D} \eta_{r,e} \int_e [\hat{u}] [\varphi] \, ds + \xi \sum_{e \in \Gamma_D} \int_e \{\varepsilon \nabla \varphi \cdot \nu\} [\hat{u}] \, ds. \end{aligned} \quad (83)$$

By adding this result side-by-side to (82) we obtain (69).



We proceed to (69)  $\Rightarrow$  (68). Assume (69) is true. First we recover the Dirichlet boundary conditions. Take any  $e \in \Gamma_D$ , such that  $e \in \partial\Omega_i$ , and  $\bar{\varphi} \in C_0^\infty(e)$ . Then let  $\{\varphi_\epsilon\}_\epsilon$  be a sequence of functions, such that

$$\begin{aligned} \varphi_\epsilon &\in C^\infty(\Omega), & \varphi_\epsilon|_e &= \bar{\varphi}, & \text{supp}(\varphi_\epsilon) &\subset \Omega_i \cup e, \\ \varphi_\epsilon|_{\partial\Omega_i \setminus e} &\equiv 0, & \nabla \varphi_\epsilon \cdot \nu \Big|_{\partial\Omega_i} &= 0, & \|\varphi_\epsilon\|_{L_2(\Omega)} &\xrightarrow{\epsilon \rightarrow 0} 0. \end{aligned}$$

Then  $\varphi \in H^1(\mathcal{E}) \cap H^2(\mathcal{T}_h)$  and (69) becomes

$$\int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi_\epsilon \, dx - \int_e \varepsilon \nabla u \cdot \nu \bar{\varphi} \, ds + \eta_{r,e} \int_e u \bar{\varphi} \, ds = \int_{\Omega_i} f \varphi_\epsilon \, dx + \eta_{r,e} \int_e \hat{u} \bar{\varphi} \, ds. \quad (84)$$

By the Green theorem

$$\int_{\Omega_i} \nabla \left( \varepsilon \nabla u \right) \varphi_\epsilon \, dx + \eta_{r,e} \int_e u \bar{\varphi} \, ds = \int_{\Omega_i} f \varphi_\epsilon \, dx + \eta_{r,e} \int_e \hat{u} \bar{\varphi} \, ds. \quad (85)$$

Passing to the limit  $\epsilon \rightarrow 0$

$$\eta_{r,e} \int_e u \bar{\varphi} \, ds = \eta_{r,e} \int_e \hat{u} \bar{\varphi} \, ds. \quad (86)$$

Since  $\bar{\varphi} \in C_0^\infty(e)$  and  $e \in \Gamma_D$  are arbitrary, we get

$$u|_{\Sigma_D} = \hat{u}|_{\Sigma_D}, \quad (87)$$

and the Dirichlet boundary conditions are satisfied.

Then take any  $\varphi \in C_{0,\Sigma_D}^\infty(\bar{\Omega})$ . Thus

$$\sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u][\varphi] \, ds = \sum_{e \in \Gamma_D} \eta_{r,e} \int_e [\hat{u}][\varphi] \, ds = 0. \quad (88)$$

as  $[\varphi] = 0$  for any  $e \in \Gamma_I$  since  $\varphi \in C_{0,\Sigma_D}^\infty(\bar{\Omega})$  and on  $e \in \Gamma_D$  we have  $[\varphi] = \varphi \equiv 0$ . Analogously we see that

$$- \sum_{e \in \Gamma_{DI}} \int_e \left\{ \varepsilon \nabla u \right\} [\varphi] \, ds = 0. \quad (89)$$

By assumptions of the theorem  $u \in H^1(\Omega)$ , so  $[u] = 0$  for any  $e \in \Gamma_I$  while as we have already been shown  $u = \hat{u}$  for  $e \in \Gamma_D$ , so

$$\xi \sum_{e \in \Gamma_{DI}} \int_e \left\{ \varepsilon \nabla \varphi \cdot \nu \right\} [u] \, ds = \xi \sum_{e \in \Gamma_D} \int_e \left\{ \varepsilon \nabla \varphi \cdot \nu \right\} [\hat{u}] \, ds. \quad (90)$$

Thus we obtain

$$\sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi \, dx = \int_{\Omega} f \varphi \, dx. \quad (91)$$

Since this statement is true for any  $\varphi \in C_{0,\Sigma_D}^\infty(\bar{\Omega})$ , then it is valid also for any  $\varphi \in H_{0,\Sigma_D}^1(\Omega)$ , so we regain the first statement of (68).  $\square$

### 7.3 Auxiliary estimates

We define the following operators

$$\begin{aligned}
A(u, \varphi) &:= \sum_{i=1}^N \int_{\Omega_i} \varepsilon \nabla u \cdot \nabla \varphi \, dx, \\
B(u, \varphi) &:= \sum_{i=1}^N \int_{\Omega_i} (e^{u-\hat{v}} - e^{\hat{w}-u}) \varphi \, dx, \\
C(\varphi) &:= \sum_{i=1}^N \int_{\Omega_i} k_1 \varphi \, dx, \\
D(u, \varphi) &:= - \sum_{e \in \Gamma_{DI}} \int_e \left\{ \varepsilon \frac{\partial u}{\partial \nu} \right\} [\varphi] \, ds, \\
E(u, \varphi) &:= - \sum_{e \in \Gamma_{DI}} \int_e \left\{ \varepsilon \frac{\partial \varphi}{\partial \nu} \right\} [u] \, ds, \\
F(\varphi) &:= - \sum_{e \in \Gamma_D} \int_e \{ \varepsilon \nabla \varphi \cdot n \} [\hat{u}] \, ds, \\
I_r(\varphi) &:= \sum_{e \in \Gamma_D} \eta_{r,e} \int_e [\hat{u}] \cdot [\varphi] \, ds, \\
J_r(u, \varphi) &:= \sum_{e \in \Gamma_{DI}} \eta_{r,e} \int_e [u] \cdot [\varphi] \, ds.
\end{aligned} \tag{92}$$

**Lemma 7.5.** *Let  $u_h \in X_h(\Omega)$  and  $\xi \in \{0, -1\}$ . Then*

$$\begin{aligned}
&A(u_h, u_h) + J(u_h, u_h) \\
&+ \xi D(u_h, u_h) + \xi E(u_h, u_h) \geq c \|u_h\|_h.
\end{aligned} \tag{93}$$

*Proof.* If  $\xi = 0$ , then simply

$$A(u_h, u_h) + J(u_h, u_h) = \|u_h\|_h^2. \tag{94}$$

Otherwise it is a simple consequence of lemma 3.3. □

**Lemma 7.6.** *Let  $u, v \in L_2(\Omega)$ . Then  $B(u, u-v) - B(v, u-v) \geq 0$ .*

*Proof.* Since the exponential function is monotone

$$\begin{aligned}
B(u, u-v) - B(v, u-v) &= \int_{\Omega} e^{-\hat{v}} (e^u - e^v) (u-v) \, dx \\
&+ \int_{\Omega} e^{\hat{w}} (e^{-v} - e^{-u}) (u-v) \, dx \geq 0.
\end{aligned} \tag{95}$$

□

**Lemma 7.7.** *Let  $u, \varphi \in H^1(\mathcal{E})$ ,  $r \in \{1, 2\}$ . Then*

$$|A(u, \varphi) + J_r(u, \varphi)| \leq C \|u\|_{h, \sigma_r} \|\varphi\|_{h, \sigma_r}. \tag{96}$$

*Proof.* This is a simple consequence of the Schwarz inequality.  $\square$

**Lemma 7.8.** *Let  $u, v, \varphi \in H^1(\mathcal{E})$ ,  $r \in \{1, 2\}$  and  $\alpha \leq u, v \leq \beta$  for some  $\alpha, \beta \in \mathbb{R}$ . Then*

$$\left| B(u, \varphi) - B(v, \varphi) \right| \leq C \|u - v\|_{h, \sigma_r} \|\varphi\|_{h, \sigma_r}, \quad (97)$$

where  $C$  is a constant dependent on  $\alpha, \beta, \|\hat{v}\|_{L_\infty(\Omega)}$  and  $\|\hat{w}\|_{L_\infty(\Omega)}$ .

*Proof.* Note that the exponential function is locally Lipschitz-continuous, so since  $u, v$  are bounded

$$\|e^u - e^v\|_{L_2(\Omega)} \leq C \|u - v\|_{L_2(\Omega)}. \quad (98)$$

The same is true for  $e^{-v} - e^{-u}$ . Thus using the Schwarz inequality and Poincare inequality for the “broken” norm (lemma 7.2)

$$\begin{aligned} \left| B(u, \varphi) - B(v, \varphi) \right| &= \left| \int_{\Omega} e^{-\hat{v}} (e^u - e^v) \varphi \, dx + \int_{\Omega} e^{\hat{w}} (e^{-v} - e^{-u}) \varphi \, dx \right| \\ &\leq C \|u - v\|_{L_2(\Omega)} \|\varphi\|_{L_2(\Omega)} \leq C \|u - v\|_{h, \sigma_r} \|\varphi\|_{h, \sigma_r}. \end{aligned} \quad (99)$$

$\square$

**Lemma 7.9.** *Let  $u \in H^2(\mathcal{E})$  and  $\varphi_h \in X_h(\Omega)$ . Then*

$$|D(u, \varphi_h)| \leq Ch |u|_{H^1(\Omega)} \|\varphi_h\|_{h, \sigma_2}. \quad (100)$$

Constant  $C$  depends on  $\varepsilon_M$  and  $\sigma_0$ .

*Proof.* Using the Schwarz inequality

$$\begin{aligned} |D(u, \varphi_h)| &\leq \sum_{e \in \Gamma_{DI}} \int_e |\{\varepsilon \nabla u \cdot \nu\} [\varphi_h]| \\ &\leq \left[ \sum_{e \in \Gamma_{DI}} \int_e \eta_{2,e}^{-1} \{\varepsilon \nabla u \cdot \nu\}^2 \right]^{1/2} \left[ \sum_{e \in \Gamma_{DI}} \int_e \eta_{2,e} [\varphi_h]^2 \right]^{1/2} \\ &\leq \left[ \sum_{i=1}^N h_i^2 \sum_{e \in \Gamma_{DI} \cap \Gamma_i} \int_e \sigma_e^{-1} \{\varepsilon \nabla u \cdot \nu\}^2 \right]^{1/2} \|\varphi_h\|_{h, \sigma_2} \\ &\leq Ch \left[ \sum_{e \in \Gamma_{DI}} \sigma_e^{-1} \{\varepsilon \nabla u \cdot \nu\}^2 \right]^{1/2} \|\varphi_h\|_{h, \sigma_2} \\ &\leq Ch |u|_{H^1(\Omega)} \|\varphi_h\|_{h, \sigma_2}, \end{aligned} \quad (101)$$

where the last inequality follows from the trace theorem. Constant  $C$  depends on  $\varepsilon_M, \sigma_0$  and the maximal number of edges of elements of the coarse triangulation  $\mathcal{E}$ .  $\square$

**Lemma 7.10.** *Let  $u \in H^2(\mathcal{E})$ ,  $u_I := \Pi_h u$  (see section 6) and  $\varphi_h \in X_h(\Omega)$ . Then*

$$|D(u - u_I, \varphi_h)| \leq Ch \left( \sum_{i=1}^N |u_i|_{H^2(\Omega_i)}^2 \right)^{1/2} \|\varphi_h\|_{h, \sigma_1}. \quad (102)$$

Constant  $C$  depends on  $\varepsilon_M$  and  $\sigma_0$ .

*Proof.* We have

$$D(u - u_I, \varphi_h) = - \sum_{e \in \Gamma_{DI}} \int_e \left\{ \varepsilon \nabla(u - u_I) \cdot \nu \right\} [\varphi_h] ds \quad (103)$$

Let us take any  $e \in \Gamma_I$ ,  $e \in \partial\Omega_j \cap \partial\Omega_k$ . Then the Schwarz inequality yields that

$$\int_e \left\{ \varepsilon \nabla(u - u_I) \cdot \nu \right\} [\varphi_h] ds \leq \varepsilon_M \|\{\nabla(u - u_I) \cdot \nu\}\|_{L_2(e)} \|\varphi_h\|_{L_2(e)} \quad (104)$$

Then by lemma 6.1

$$\begin{aligned} \|\{\nabla(u - u_I) \cdot \nu\}\|_{L_2(e)} &\leq (h_j^{1/2} |u_j|_{H^2(\Omega_j)} + h_k^{1/2} |u_k|_{H^2(\Omega_k)}) \\ &\leq (h_j + h_k)^{1/2} (|u_j|_{H^2(\Omega_j)} + |u_k|_{H^2(\Omega_k)}). \end{aligned} \quad (105)$$

Therefore

$$\begin{aligned} \eta_{1,e}^{-1} \|\{\nabla(u - u_I)\}\|_{L_2(e)}^2 &\leq C \sigma_e^{-1} h_j h_k (|u_j|_{H^2(\Omega_j)} + |u_k|_{H^2(\Omega_k)})^2 \\ &\leq C h^2 (|u_j|_{H^2(\Omega_j)} + |u_k|_{H^2(\Omega_k)})^2 \end{aligned} \quad (106)$$

If  $e \in \Gamma_D$ ,  $e \in \partial\Omega_i$ , then analogously

$$\eta_{1,e}^{-1} \|\{\nabla(u - u_I) \cdot \nu\}\|_{L_2(e)}^2 \leq C \sigma_e^{-1} h_i^2 |u_i|_{H^2(\Omega_i)}^2 \leq C h^2 |u_i|_{H^2(\Omega_i)}^2 \quad (107)$$

Therefore by Schwarz inequality and the inequalities derived above

$$\begin{aligned} &\sum_{e \in \Gamma_{DI}} \int_e \left\{ \varepsilon \nabla(u - u_I) \cdot \nu \right\} [\varphi_h] ds \\ &\leq C \left( \sum_{e \in \Gamma_{DI}} \eta_{1,e}^{-1} \|\{\nabla(u - u_I) \cdot \nu\}\|_{L_2(e)}^2 \right)^{1/2} \left( \sum_{e \in \Gamma_{DI}} \eta_{1,e} \|\varphi_h\|_{L_2(e)}^2 \right)^{1/2} \\ &\leq C h \left( \sum_{i=1}^N |u_i|_{H^2(\Omega_i)}^2 \right)^{1/2} \|\varphi_h\|_{h, \sigma_1}. \end{aligned} \quad (108)$$

Constant  $C$  is independent of  $h$ . It depends on  $\sigma_0$ ,  $\varepsilon_M$  and on the number of elements of  $\Gamma_{DI}$ .  $\square$

**Lemma 7.11.** *Let  $u \in H^2(\mathcal{E})$ ,  $u_I := \Pi_h u$  (see section 6) and  $\varphi_h \in X_h(\Omega)$ . Then*

$$|E(u - u_I, \varphi_h)| \leq C \|\varphi_h\|_{h, \sigma_1} \left[ \sum_{i=1}^N \left( h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k} \right) |u_i|_{H^2(\Omega_i)}^2 \right]^{1/2}. \quad (109)$$

*Proof.* By the Schwarz inequality

$$\begin{aligned} |E(u - u_I, \varphi_h)| &\leq \sum_{e \in \Gamma_{DI}} \int_e \left| \left\{ \varepsilon \nabla \varphi_h \cdot \nu \right\} \right| |u - u_I| ds, \\ &\leq \varepsilon_M \sum_{e \in \Gamma_{DI}} \|\{\nabla \varphi_h \cdot \nu\}\|_{L_2(e)} \|u - u_I\|_{L_2(e)} \end{aligned} \quad (110)$$

Splitting this sum up

$$\|\{\nabla\varphi_h \cdot \nu\}\|_{L_2(e)} \leq \left\| \nabla\varphi_h \cdot \nu \Big|_{\Omega_j} \right\|_{L_2(e)} + \left\| \nabla\varphi_h \cdot \nu \Big|_{\Omega_k} \right\|_{L_2(e)}. \quad (111)$$

Then using lemma 3.1

$$\left\| \nabla\varphi_h \cdot \nu \Big|_{\Omega_i} \right\|_{L_2(e)}^2 \leq Ch_i^{-1} \|\nabla\varphi_h\|_{L_2(\Omega_i)}^2 \leq Ch_i^{-1} \|\varphi_h\|_{h,\sigma_1}^2. \quad (112)$$

On the other hand

$$\|u - u_I\|_{L_2(e)} \leq \|u - u_I\|_{\Omega_j} + \|u - u_I\|_{\Omega_k}. \quad (113)$$

By lemma 6.1 we have  $\|u - u_I\|_{\Omega_i}^2 \leq Ch_i^3 |u|_{H^2(\Omega_i)}^2$ , so

$$|E(u - u_I, \varphi_h)| \leq C \|\varphi_h\|_{h,\sigma_1} \left[ \sum_{i=1}^N \left( h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k} \right) |u_i|_{H^2(\Omega_i)}^2 \right]^{1/2} \quad (114)$$

□

The differential problem (69) satisfies:

$$\begin{aligned} A(u^*, \varphi) + B(u^*, \varphi) + D(u^*, \varphi) + \xi E(u^*, \varphi) + J_r(u^*, \varphi) \\ = C(\varphi) + \xi F(\varphi) + I_r(\varphi) \quad \forall \varphi \in H^1(\mathcal{E}) \cap H^2(\mathcal{T}_h). \end{aligned} \quad (115)$$

On the other hand, the family of discrete problems depending on parameter  $h$  is defined as

$$\begin{aligned} A(u_h^*, \varphi_h) + B(u_h^*, \varphi_h) + \xi D(u_h^*, \varphi_h) + \xi E(u_h^*, \varphi_h) + J_r(u_h^*, \varphi_h) \\ = C(\varphi_h) + \xi F(\varphi_h) + I_r(\varphi_h) \quad \forall \varphi_h \in X_h. \end{aligned} \quad (116)$$

Here we use  $\xi = 0$  for CWOPSIP (16) and  $\xi = 1$  for CSIPG (18).

We subtract these equations from each other with  $\varphi := \varphi_h$  and we obtain

$$\begin{aligned} A(u^* - u_h^*, \varphi_h) + B(u^*, \varphi_h) - B(u_h^*, \varphi_h) + D(u^* - \xi u_h^*, \varphi_h) \\ + \xi E(u^* - u_h^*, \varphi_h) + J_r(u^* - u_h^*, \varphi_h) = 0. \end{aligned} \quad (117)$$

This is equivalent to LHS = RHS, where

$$\begin{aligned} \text{LHS} := A(u_I^* - u_h^*, \varphi_h) + B(u_I^*, \varphi_h) - B(u_h^*, \varphi_h) + \xi D(u_I^* - u_h^*, \varphi_h) \\ + \xi E(u_I^* - u_h^*, \varphi_h) + J_r(u_I^* - u_h^*, \varphi_h), \end{aligned} \quad (118)$$

and

$$\begin{aligned} \text{RHS} := A(u_I^* - u^*, \varphi_h) + B(u_I^*, \varphi_h) - B(u^*, \varphi_h) + D(\xi u_I^* - u^*, \varphi_h) \\ + \xi E(u_I^* - u^*, \varphi_h) + J_r(u_I^* - u^*, \varphi_h). \end{aligned} \quad (119)$$

## 7.4 Error estimate for CWOPSIP

In this case

$$\text{RHS} := A(u_I^* - u^*, \varphi_h) + B(u_I^*, \varphi_h) - B(u^*, \varphi_h) - D(u^*, \varphi_h) + J_2(u_I^* - u^*, \varphi_h). \quad (120)$$

We take  $\varphi_h := u_I^* - u_h^*$ . By lemma 7.5 and lemma 7.6 we have that  $\text{LHS} \geq c\|u_I^* - u_h^*\|_{h,\sigma_2}$ . Note that  $u^*$  is bounded (lemma 2.1), thus  $u_I^*$  is also bounded by the same constants, what allows us to use lemma 7.6.

On the other hand, we may estimate RHS with lemmata 7.7, 7.8 and 7.9

$$\text{RHS} \leq C\|u_I^* - u_h^*\|_{h,\sigma_2} \left( \|u^* - u_I^*\|_{h,\sigma_2} + h|u^*|_{H^1(\Omega)} \right). \quad (121)$$

Therefore using  $\text{LHS} = \text{RHS}$  we obtain

$$\|u_I^* - u_h^*\|_{h,\sigma_2}^2 \leq C\|u_I^* - u_h^*\|_{h,\sigma_2} \left( \|u^* - u_I^*\|_{h,\sigma_2} + h|u^*|_{H^1(\Omega)} \right). \quad (122)$$

Then diving both sides of this inequality by  $\|u_I^* - u_h^*\|_{h,\sigma_2} > 0$

$$\|u_I^* - u_h^*\|_{h,\sigma_2} \leq C \left( \|u^* - u_I^*\|_{h,\sigma_2} + h|u^*|_{H^1(\Omega)} \right). \quad (123)$$

Thus by the triangle inequality and interpolation error estimate (60) we have

$$\begin{aligned} \|u^* - u_h^*\|_{h,\sigma_2} &\leq \|u^* - u_I^*\|_{h,\sigma_2} + \|u_I^* - u_h^*\|_{h,\sigma_2} \\ &\leq C \left( \sum_{i=1}^N \left( h_i + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k^2} \right) |u_i^*|_{H^2(\Omega_i)}^2 \right)^{1/2} + Ch|u^*|_{H^1(\Omega)}. \end{aligned} \quad (124)$$

With the additional assumption  $h_i := c_i h$  for every  $\Omega_i \in \mathcal{E}$  we can simplify this expression to

$$\|u^* - u_h^*\|_{h,\sigma_2} \leq Ch^{1/2} \left( |u^*|_{H^1(\Omega)}^2 + \sum_{i=1}^N |u_i^*|_{H^2(\Omega_i)}^2 \right)^{1/2}. \quad (125)$$

If we use particular grids and Dirichlet boundary conditions are piecewise linear, as explained in section 6, we could get a better estimate

$$\|u^* - u_h^*\|_{h,\sigma_2} \leq Ch \left( |u^*|_{H^1(\Omega)}^2 + \sum_{i=1}^N |u_i^*|_{H^2(\Omega_i)}^2 \right)^{1/2}. \quad (126)$$

## 7.5 Error estimate for CSIPG

As for CWOPSIP, by lemma 7.5 and lemma 7.6 imply  $\text{LHS} \geq c\|u_I^* - u_h^*\|_{h,\sigma_1}$ . Then we have

$$\begin{aligned} \text{RHS} &:= A(u_I^* - u^*, \varphi_h) + B(u_I^*, \varphi_h) - B(u^*, \varphi_h) + D(u_I^* - u^*, \varphi_h) \\ &\quad + E(u_I^* - u^*, \varphi_h) + J_1(u_I^* - u^*, \varphi_h). \end{aligned} \quad (127)$$

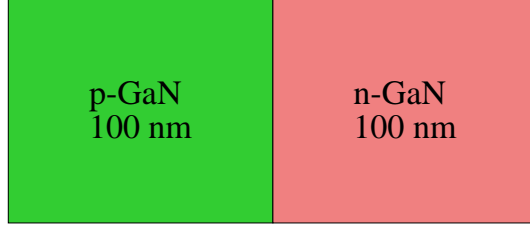


Figure 4: Schema of the n-p diode used in numerical simulations.

This time we may estimate RHS with lemmata 7.7, 7.8, 7.10 and 7.11

$$\text{RHS} \leq C \|u_I^* - u_h^*\|_{h, \sigma_1} \left( \|u_I^* - u^*\|_{h, \sigma_1} + \left[ \sum_{i=1}^N \left( h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k} \right) |u_i^*|_{H^2(\Omega_i)}^2 \right]^{1/2} \right). \quad (128)$$

Thus estimating LHS = RHS from below and above and dividing by  $\|u_I^* - u^*\|_{h, \sigma_1} > 0$  we obtain

$$\|u_I^* - u_h^*\|_{h, \sigma_1} \leq \left( \|u_I^* - u^*\|_{h, \sigma_1} + \left[ \sum_{i=1}^N \left( h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k} \right) |u_i^*|_{H^2(\Omega_i)}^2 \right]^{1/2} \right). \quad (129)$$

Thus by the triangle inequality and interpolation error estimate (56) we have

$$\begin{aligned} \|u^* - u_h^*\|_{h, \sigma_1} &\leq \|u^* - u_I^*\|_{h, \sigma_1} + \|u_I^* - u_h^*\|_{h, \sigma_1} \\ &\leq C \left( \sum_{i=1}^N \left( h_i^2 + \sum_{\Omega_k \in \text{nb}(\Omega_i)} \frac{h_i^3}{h_k} \right) |u_i^*|_{H^2(\Omega_i)}^2 \right)^{1/2}. \end{aligned} \quad (130)$$

If  $h_i := c_i h$  for every  $\Omega_i \in \mathcal{E}$  this estimate simplifies to

$$\|u^* - u_h^*\|_{h, \sigma_1} \leq C h \left( \sum_{i=1}^N |u_i^*|_{H^2(\Omega_i)}^2 \right)^{1/2}. \quad (131)$$

## 8 Numerical experiments

We performed numerical simulations of a p-n junction (figure 4) to verify whether the theoretical estimates hold in practice. This device consists of two layers composed of the same material, but with different doping. We use homogeneous mesh inside layers of the device, which is nonmatching on the interface of the layers. It is formed by division of the layers to  $K$  parts in the longitude and then by dividing the first layer into  $3K$  parts and dividing the second part into  $2K$  parts in the perpendicular direction. We start with initial grid for  $K = 2$ , presented in figure 5.

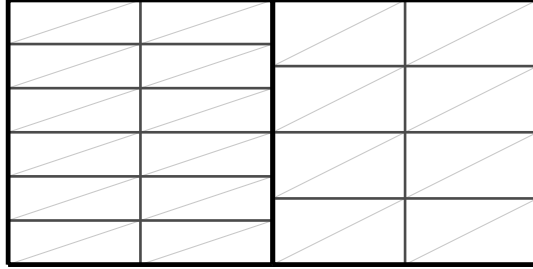


Figure 5: Initial grid used in numerical simulations of pn junction.

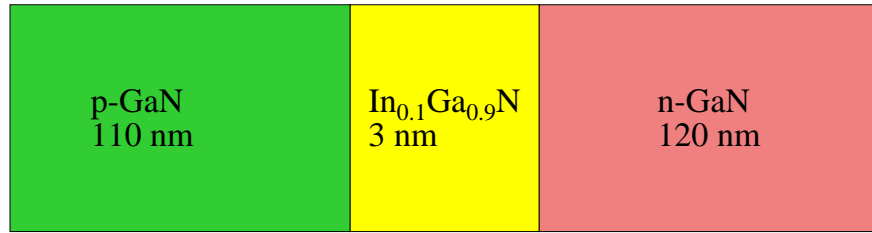


Figure 6: Schema of the quantum well structure used in numerical simulations.

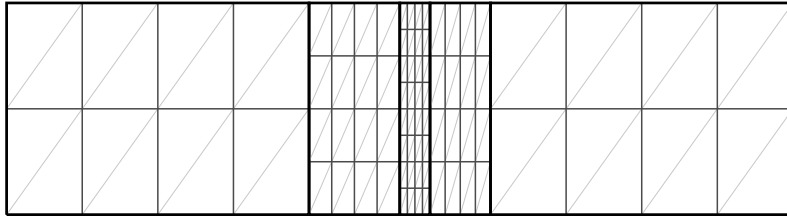


Figure 7: Initial grid used in numerical simulations of the quantum well. To improve the convergence, p-GaN and n-GaN were splitted into two layers to thicken the grid near the quantum well.



Table 1: Relative errors of the potential  $u^*$ . Simulation were performed for the pn junction in the equilibrium state.

K	CSIPG		CWOPSIP	
	$L_2(\Omega)$	$H^1(\Omega)$	$L_2(\Omega)$	$H^1(\Omega)$
2	1.8e-01	6.9e-01	2.3e-01	9.8e-01
4	4.5e-02 (4.0)	4.2e-01 (1.6)	1.2e-01 (1.9)	5.9e-01 (1.7)
8	2.3e-02 (1.9)	3.0e-01 (1.4)	3.8e-02 (3.1)	3.2e-01 (1.9)
16	9.7e-03 (2.4)	2.1e-01 (1.4)	7.7e-03 (4.9)	2.1e-01 (1.5)
32	2.8e-03 (3.4)	1.2e-01 (1.8)	1.5e-03 (5.1)	1.2e-01 (1.8)
64	6.1e-04 (4.6)	5.7e-02 (2.0)	4.8e-04 (3.1)	5.7e-02 (2.0)
128	1.5e-04 (4.2)	2.9e-02 (2.0)	1.3e-04 (3.8)	2.9e-02 (2.0)
256	3.5e-05 (4.2)	1.4e-02 (2.0)	3.3e-05 (3.9)	1.4e-02 (2.0)

In numerical simulations, we use the Poisson equation scaled in SI units. Thus it is more elaborated than equation (1), as then physical constants and material parameters are present. Also it takes into account physical parameters and incomplete ionization of acceptors and donors, so  $k_1$  is actually a function of  $u$ . More details on the drift-diffusion equations, used by us in the simulations of realistic devices may be found in [16], while details of the linearization of the underlying discrete system may be found in [15]. Since the exact solution is not known, as a reference we use the result of one-dimensional simulation for  $K = 1024$ . For every function  $f_K$  taken into account, we compute the relative errors defined as

$$\text{error}_{K,L_2(\Omega)} := \frac{\|f_K - f_{\text{ref}}\|_{L_2(\Omega)}}{\|f_{\text{ref}}\|_{L_2(\Omega)}}, \quad \text{error}_{K,H^1(\Omega)} := \frac{\|f_K - f_{\text{ref}}\|_{H^1(\Omega)}}{\|f_{\text{ref}}\|_{H^1(\Omega)}}, \quad (132)$$

where  $f_{\text{ref}}$  is a numerical solution computed on a fine grid, as mentioned before. Errors are presented in function of  $K$ . Note that  $h = c/K$  for some constant  $c$ . Having in mind estimates of section 7, we expect  $\text{error}_{K,H^1(\Omega)} / \text{error}_{2K,H^1(\Omega)} \approx 2$ .

Results of these simulations are presented in table 1. Both methods start slowly, for  $K \leq 32$  error ratio is  $< 2$ . From  $K = 64$  we observe error ratio  $\approx 2$ . The error norms for CSIPG and CWOPSIP are similar.

Then we repeated our simulations for a single quantum well structure (figure 6). It is similar to the p-n junction, but it has a narrow layer between n region and p region, called the quantum well. In this case, we introduce five layers in our mesh, while two additional layers are used to improve the grid in n region and p region near the quantum well layer (figure 7). Results of this simulation is presented in table 2 and they generally agree with our previous simulation. Note that the slow start is absent in this case. This is due to additional layers introduced near the quantum well, where the function fluctuations are crucial.

## 9 Conclusions

We have presented two methods of composite Discontinuous Galerkin discretization of the drift-diffusion equations in the equilibrium state. These methods were derived from Symmetric Interior Penalty Galerkin method [13] and Weakly Over-Penalized Symmetric Interior Penalty method [3].

Table 2: Relative errors of the potential  $u^*$ . Simulation were performed for the single quantum well in the equilibrium state.

K	CSIPG		CWOPSIP	
	$L_2(\Omega)$	$H^1(\Omega)$	$L_2(\Omega)$	$H^1(\Omega)$
2	5.2e-02	3.8e-01	7.5e-02	4.6e-01
4	2.0e-02 (2.6)	2.3e-01 (1.7)	2.4e-02 (3.1)	2.6e-01 (1.8)
8	5.2e-03 (3.8)	1.1e-01 (2.0)	7.3e-03 (3.3)	1.2e-01 (2.2)
16	1.3e-03 (3.9)	5.8e-02 (1.9)	1.8e-03 (4.0)	5.9e-02 (2.0)
32	3.4e-04 (4.0)	2.9e-02 (2.0)	4.5e-04 (4.0)	3.0e-02 (2.0)
64	8.5e-05 (4.0)	1.5e-02 (2.0)	1.1e-04 (4.0)	1.5e-02 (2.0)

Both discretizations are shown to be well-defined and the error is estimated. In case of uniform increase of grid density, the  $H^1$ -norm of error of Composite Symmetric Interior Penalty Galerkin (CSIPG) method error is estimated by  $O(h)$ , while for Composite Weakly Over-Penalized Symmetric Interior Penalty (WOPSIP) method we obtain only  $O(h^{1/2})$  estimate. In the latter case, the estimate may be improved to  $O(h)$  under additional assumptions on a mesh sequence and boundary conditions. In simulations of semiconductor luminescent devices, these assumptions are not excessively restrictive, as these structures' design is often based on simple geometric shapes.

Numerical simulations presented in this paper stay in a good agreement with these estimates.

## References

- [1] Randolph E. Bank and Donald J. Rose. Some Error Estimates for the Box Method. *SIAM Journal of Numerical Analysis*, 24:777–787, 1987.
- [2] A. T. Barker, S. C. Brenner, E.-H. Park, and L.-Y. Sung. Two-Level Additive Schwarz Preconditioners for a Weakly Over-Penalized Symmetric Interior Penalty Method. *Journal of Scientific Computing*, 47(1):27, 2011.
- [3] S. C. Brenner, L. Owens, and L.-Y. Sung. A weakly over-penalized symmetric interior penalty method. *Electronic Transactions on Numerical Analysis*, 30:107–127, 2008.
- [4] Susanne C. Brenner. Poincare-Friedrichs Inequalities for Piecewise  $H^1$  Functions. *SIAM Journal on Numerical Analysis*, 41:306–324, 2004.
- [5] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, 2008.
- [6] R. K. Coomer and I. G. Graham. Massively Parallel Methods for Semiconductor Device Modelling. *Computing*, 56:1–27, 1996.
- [7] Maksymilian Dryja. On Discontinuous Galerkin Methods for Elliptic Problems with Discontinuous Coefficients. *Computational Methods in Applied Mathematics*, 3(1):76–85, 2003.
- [8] H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *IEEE Trans. Elect. Dev.*, 11:455–464, 1964.

- [9] Joseph W. Jerome. Consistency of Semiconductor Modeling: An Existence/Stability Analysis for the Stationary van Roosbroeck System. *SIAM Journal of Applied Mathematics*, 45(4):565–590, 1985.
- [10] J. L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod/Gauthier-Villars, Paris, 1969.
- [11] J. J. H. Miller, W. H. A. Schilders, and S. Wang. Application of finite element methods to the simulation of semiconductor devices. *Rep. Prog. Phys.*, 62:277–353, 1999.
- [12] Daniele A. Di Pietro and Alexandre Ern. *Mathematical aspects of Discontinuous Galerkin Methods*. Springer, Berlin, 2012.
- [13] Beatrice Riviere. *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation*. Society for Industrial and Applied Mathematics, Philadelphia, 2008.
- [14] W. V. Van Roosbroeck. Theory of Flow of Electrons and Holes in Germanium and Other Semiconductors. *The Bell System Technical Journal*, 29:560–607, 1950.
- [15] Konrad Sakowski, Leszek Marcinkowski, and Stanislaw Krukowski. Modification of the Newton’s method for the simulations of gallium nitride semiconductor devices. *Lecture Notes in Computer Science*, 8385:551–560, 2014.
- [16] Konrad Sakowski, Leszek Marcinkowski, Stanislaw Krukowski, Szymon Grzanka, and Elzbieta Litwin-Staszewska. Simulation of trap-assisted tunneling effect on characteristics of gallium nitride diodes. *Journal of Applied Physics*, 111(12):123115, 2012.
- [17] Siegfried Selberherr. *Analysis and Simulation of Semiconductor Devices*. Springer-Verlag, Wien, 1984.
- [18] S.M. Sze and K.K. Ng. *Physics of Semiconductor Devices*. Wiley-Interscience, Berlin, 2006.
- [19] Peter Wilkes. *Solid State Theory in Metallurgy*. Cambridge University Press, Cambridge, 1973.